

2019 가을학기

데이터 저널리즘

담당교수 이준환

email: joonhwan@snu.ac.kr

cell: 010-9212-4975

office: 64동 405호

시간 월요일 오후 3:00 ~ 5:50**장소** 83-601

수업 개요 데이터 저널리즘은 디지털 사회에서의 정보의 생산과 확산 과정에서 점차 중요해지
는 데이터의 역할을 다루는 학문 분야이다. 본 과목에서는 데이터 저널리즘의 이해와
더불어 이를 수행할 수 있는 컴퓨팅 기술을 학습하는 것을 목적으로 한다. 이를 위해
1) 데이터 저널리즘의 개념과 사례를 조사, 학습하고 2) 파이썬 프로그래밍 언어를 통
한 데이터 분석 기법을 학습하고 이를 데이터 저널리즘에 활용하는 방법을 다룬다.

또한 본 수업에서는 팀별 프로젝트를 통해 실전적인 데이터 저널리즘에 대한 이해와
해당 분야에 대한 포괄적 경험을 구축하고자 한다.

- 강의내용**
1. Python 을 이용한 기초 프로그래밍
 2. Python 기반의 라이브러리를 이용한 데이터 처리 프로그래밍
 3. 웹 기반 기술(web technology)
 4. 소셜 데이터 마이닝(social data mining)을 통한 데이터 분석

- 교재**
- [CodeCademy: Learn Python](#)
 - [CodeCademy: Learn HTML & CSS](#)
 - [파이썬 라이브러리를 활용한 데이터 분석 \(Python for Data Analysis\)](#)
 - [데이터 저널리즘: 스토리텔링의 과학 \(The Art and Science of Data-Driven Journalism\)](#)
 - 기타 필요한 교재/논문 등은 강의 중 제공

수업진행 계획 Module 1: Data Science for Journalist

- 1주 **What is Data Journalism?**
 - 데이터 저널리즘의 개념 및 사례 소개
Tools for Data Journalism (Python, R 등)
 - 왜 Python 인가?
 - Python의 설치
 - Jupyter Notebook, numpy, pandas 등의 설치
 - Editor의 선택
- 2주 **Python Crash Course 1**
 - 기초 Python 문법 학습
 - Github 사용법
- 3주 **Python Crash Course 2**
 - 기초 Python 문법 학습
- 4주 **Data Processing**
 - Data cleaning process
- 5주 **Data Analysis Using numpy and pandas**
 - numpy와 pandas를 활용한 데이터의 분석
- 6주 **Text Data Processing**
 - NLTK를 활용한 텍스트의 처리
 - KoNLPy를 활용한 한글 텍스트의 처리
- 7주 **Data Visualization**
 - 데이터 시각화 방법론과 도구의 소개
 - matplotlib, ggplot2, plotly의 학습
- 8주 **Web Technology Basics**
 - Internet Basis
 - HTML & CSS
 - HTML 5
- 9주 **Social Data Mining 1**
 - Crawling data from Website
 - Crawling data from Twitter
- 10주 **Social Data Mining 2**
 - Crawling data from Social Media
- 11주 **Robot Journalism**
 - 로봇저널리즘의 이해와 실습
- 12주 **Final Exam**
- Module 2: Final Project
- 13주 **Project idea presentation**
- 14주 **Team Meeting**
- 15주 **Final Presentation**

- 과제 및 평가** Assignment 1: 2 CodeCademy Python 과정 수강 (5+5 points)
(Python & HTML/CSS)
- Assignment 2: 파이썬 코딩 연습 (10 points)
- Assignment 3: 데이터 클리닝 (10 points)
- Assignment 4: 텍스트 처리 (10 points)
- Assignment 5: 웹상의 데이터를 수집한 후 통계 처리 및 시각화 (30 points)
- Assignment 6: 로봇 저널리즘 - 날씨 기사 생성 (30 points)
-
- Team Project: Final Project (The Upshot 스타일의 보고서) (50 points)
- Peer Review: 30 points
- if (absent >=5): fail